МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

КРИПТОГРАФІЯ

Комп’ютерний практикум

Робота №1

Виконали:
Сернова А.Р., Колесник А.М.
студенти групи ФБ-93

Перевірила:
Селюх П.В.

Київ-2021

Мета роботи: Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела

Постановка задачі: реалізувати програму, що підраховує значення ентропії для літер та біграм (з перетином та без) у текстовому файлі, як з урахуванням пробілів, так і без них.

Під час виконання лаб. роботи виникли деякі труднощі: найбільша це вибір іншої мови програмування (вона була обрана через те, що python – дуже багата на функції з рядками мова), також деякі труднощі з фільтруванням тексту, які ми вирішили за допомогою різних інтернет ресурсів таких як stackoverflow та подібних. Дещо виникали складності під час початку роботи з біграмами, а саме продумування варіантів та алгоритмів їх підрахунку, які були вирішені розбиттям тексту на біграми і використання функції count.

Нижче наведені таблиці для літер російського алфавіту без «ё» і «ъ» без урахування пробілів (*таб. 1*), та з пробілами (*таб. 2*)

о	0.1101
е	0.0841
а	0.0790
т	0.0661
и	0.0657
н	0.0629
с	0.0552
л	0.0454
р	0.0430
в	0.0410
к	0.0388
м	0.0319
у	0.0292
п	0.0288
д	0.0287
я	0.0240
ы	0.0196
ь	0.0193
г	0.0176
ч	0.0173
з	0.0168
б	0.0163
й	0.0104
ж	0.0100
х	0.0094
ш	0.0091
ю	0.0069
щ	0.0041
э	0.0034

_	0.1644
о	0.0920
е	0.0703
а	0.0660
т	0.0552
и	0.0549
н	0.0526
с	0.0462
л	0.0378
р	0.0359
в	0.0343
к	0.0324
м	0.0267
у	0.0244
п	0.02404
д	0.02402
я	0.0200
ы	0.0163
ь	0.0162
г	0.0147
ч	0.0145
з	0.0141
б	0.0136
й	0.0087
ж	0.0084
х	0.0079
ш	0.0076
ю	0.0058
щ	0.0035

ц	0.0028
ф	0.0022

Таб. 1

э	0.0028
ц	0.0023
ф	0.0018

Таб. 2

H = 4.475

H = 4.384

Для біграм без урахування пробілів наведено фрагменти з отриманого файлу after.txt: з перетинами (*таб. 3*) та без (*таб. 4*)

```
Bigram cross, no spaces
аа is 0.0002151419259098 64 аб is 0.0016787846344620096 ав is 0.004372293785617 аг is 0.0010909559077634048 ад is 0.0027756696504197807 ае is 0.004371446770160662 аж is
0.0016677734335529615 аз is 0.0041986556170677 01 аи is 0.0010723215677239679 ай is 0.0007292803079070587 ак is 0.0073580232692 08616 ал is 0.009340039437039647 ам is
0.004755144771881797 ан is 0.004959275496859266 ао is 0.0011316122649667631 ап is 0.0026054195436958335 ар is 0.0027129905066507654 ас is 0.006131544888431124 ат is
0.006256056160512817 ау is 0.0005556421393577 59 аф is 0.0001854963849380 3235 ах is 0.0014416203066873565 ац is 0.0001592389057915529 ач is 0.0020472363579690602 аш is
0.001200220901631013 ащ is 0.0004277428054507139 аы is 0.0 аь is 0.0 аэ is 0.0002591867296394424 6 ао is 0.0024368634678845 62 ая is 0.0029018749534141497
ба is 0.0015602024705746 83 бб is 2.1175386408451182e-05 бв is 4.150375736056432e-05 бг is 2.2022401864789228e-05 бд is 1.6940309126760946e-05 бе is 0.002186993908264838 бж is
1.4399262757746803e-05 бз is 2.2022401864789228e-05 би is 0.0010646984286169254 бй is 0.0 бк is 0.0001702501067239475 бл is 0.0008309221626676244 бм is 4.997391192394479e-05 бн
is 0.0003252539352338 1017 бо is 0.0022445909592958254 бп is 1.1011200932394614e-05 бр is 0.0012213962880 39464 бс is 0.0002422464205126815 бт is 1.4399262757746803e-05 бу is
0.0013408254673831289 бф is 0.0 бх is 2.7104494602817514e-05 бц is 7.623139107042425e-06 бч is 9.31717001971852e-06 бш is 7.623139107042425e-06 бщ is 0.0003709927698760647 бы is
0.0033745095780507803 бь is 0.0002608807605521186 бэ is 0.0001143470866056363 8 бю is 5.929108194366331e-06 бя is 0.0007487616634028338
ва is 0.0069692431747494 53 вб is 0.0001507687512281724 вв is 0.0003421942443605711 вг is 0.0002473285132507098 вд is 0.0006911646123718466 ве is 0.0049228538322367305 вж is
7.1149298332395 97e-05 вз is 0.0006733772877887475 ви is 0.0031966363322197904 вй is 0.0 вк is 0.0009435752183605846 вл is 0.0010155715321493186 вм is 0.0003625226153126842 вн is
0.0016660794026 16939 во is 0.0074198553975212 94 вп is 0.0008673438272901604 вр is 0.0010469111040338265 вс is 0.0037963232753071277 вт is 0.0007690900343549469 ву is
0.0008326161935803004 вф is 3.049255642816 97e-05 вх is 8.131348380845254e-05 вц is 7.70784065267623e-05 вч is 0.0002795151005915556 вш is 0.0009384931256225564 вщ is
5.929108194366331e-06 вы is 0.0028679943351 60628 вь is 0.0001278993339070 4514 вэ is 0.0001838023540253 5626 вю is 3.3880618253521892e-06 вя is 0.0004472241604648896
га is 0.0014467023994253847 гб is 5.3361973749296974e-05 гв is 9.063065382817105e-05 гг is 8.470154563380473e-06 гд is 0.0014365382 13949328 ге is 0.0001719441376366 2358 гж is
1.6940309126760946e-06 гз is 3.896271099155017e-05 ги is 0.0010299707949070654 гй is 0.0 гк is 0.0001211232102563 4075 гл is 0.0010965719590451088 гм is 2.5410463690141418e-05 гн
is 0.0003879330790028 2563 го is 0.0090478191046030 21 гп is 7.114929833239597e-05 гр is 0.0010376469185577698 гс is 8.554856109014277e-05 гт is 6.776123650704378e-05 гу is
0.0007258922460817065 гф is 0.0 гх is 5.929108194366331e-06 гц is 3.3880618253521892e-06 гч is 6.776123650704378e-05 гш is 5.929108194366331e-06 гщ is 1.6940309126760946e-06 гы
is 0.0 гь is 0.0 гэ is 9.317170019718 52e-06 гю is 0.0 гя is 1.2705231845070709e-05
да is 0.0046509618707522 17 дб is 0.0001126530556929 6020 дв is 0.0012324074889718587 дг is 6.522019013802965e-05 дд is 9.8253792935213 48e-05 де is 0.004766155972814192 дж is
9.317170019718519e-05 дз is 5.166794283662088e-05 ди is 0.0027587293412 93302 дй is 0.0 дк is 0.0004040263726 7324853 дл is 0.0007724780961802991 дм is 0.0001236642566253549 дн is
0.001975240044180326 до is 0.0048271410856705315 дп is 0.0002981494406309926 др is 0.0013213441118873537 дс is 0.0004692465628112782 дт is 0.0003371121516225428 ду is
0.0020116617088028624 дф is 5.082092738028284e-06 дх is 7.70784065267623e-05 дц is 0.0001338284421014148 дч is 8.300751472112863e-05 дш is 0.0001084179784112700 6 дщ is
1.6940309126760946e-06 ды is 0.0008681908427464985 дь is 0.0005598772166394492 дэ is 1.3552247301408757e-05 дю is 1.3552247301408757e-05 дя is 0.000520914505647899
еа is 0.0002168359568225401 еб is 0.0024851433488958306 ев is 0.0034490469382085285 ег is 0.0044476781612 3086 ед is 0.0040173943094113 58 ее is 0.0019328892713634238 еж is
0.0010596163358788972 ез is 0.0023674082004 64842 еи is 0.0011011200932394614 ей is 0.0019328892713634238 ек is 0.0023784194013972367 ел is 0.005855417849664921 ем is
0.005630111738279 ен is 0.0085734904493 5714 ео is 0.0016542211862282062 еп is 0.003993677876633893 ер is 0.007767131734619894 ес is 0.0068667543045325494 ет is
0.0104504767002 98828 еу is 0.0008393923172310048 еф is 8.554856109014277e-05 ех is 0.0009452692492732608 ец is 0.0003836980017211354 еч is 0.0023716432777465325 еш is
0.0014873591411329611 ещ is 0.0009317170019 71852 еы is 0.0 еь is 0.0 еэ is 0.0003150897497 5775355 ею is 0.0001422985966647 9194 ея is 0.0005158324129098708
жа is 0.0016034002588479234 жб is 1.4399262757746803e-05 жв is 1.6093293670422896e-05 жг is 2.795151005915556e-05 жд is 0.0009325640174281901 же is 0.0036260731685831805 жж is
1.2705231845070709e-05 жз is 5.929108194366331e-06 жи is 0.0017312995927549685 жй is 0.0 жк is 0.0001965075858704269 7 жл is 2.8798525515493607e-05 жм is 3.5574649166197985e-05
жн is 0.0010985790468704473 жо is 9.655976202253739e-05 жп is 2.3716432777465325e-05 жр is 2.0328370952113135e-05 жс is 1.9481355495775086e-05 жт is 1.6093293670422896e-05 жу is
0.0004556943155098694 жф is 0.0 жх is 8.470154563380473e-07 жц is 0.0 жч is 2.2869417321127275e-05 жш is 0.0 жщ is 0.0 жы is 0.0 жь is 1.863434003943704e-05 жэ is
```

таб. 3

H = 7.2171

```
Bigram for no cross, no spaces
аа is 0.0002015896786 0845525 аб is 0.0016889488199380662 ав is 0.004336719136450802 аг is 0.0010452170731211504 ад is 0.002778210696788795 ае is 0.004438360991211368 аж is
0.0016177995216056702 аз is 0.0041486817051437554 аи is 0.0010977320314141093 ай is 0.0007064108905859314 ак is 0.007463900201250873 ал is 0.009291759556028378 ам is
0.004731428339104332 ан is 0.0050092494087832 12 ао is 0.0010587693204225591 ап is 0.0025596807090535 79 ар is 0.0026545464401634402 ас is 0.006113757563848025 ат is
0.0063187353042818 32 ау is 0.0005336197374929697 аф is 0.0001863434003943703 8 ах is 0.0014297620902986238 ац is 0.0001761792149183138 3 ач is 0.0020785759298535 68 аш is
0.0012247843498648164 ащ is 0.000394709202652 53 аы is 0.0 аь is 0.0 аэ is 0.0002659628532 9014683 ао is 0.0024715911015944 22 ая is 0.0029594720044451373
ба is 0.0015551203778366547 66 is 2.5410463690141418e-05 бв is 3.726868007887408e-05 бг is 1.6940309126760946e-05 бд is 2.0328370952113135e-05 бе is 0.0021056804244563854 бж is
2.2022401864789228e-05 бз is 1.6940309126760946e-05 би is 0.0010604633513352352 бй is 0.0 бк is 0.0001795672767436 66 бл is 0.0007758661580056513 бм is 4.4044803729578457e-05 бн
is 0.0003049255642816 97 бо is 0.0022327327429070925 бп is 1.863434003943704e-05 бр is 0.0012451127208165294 бс is 0.0002507165750760 6197 бт is 1.863434003943704e-05 бу is
0.0013569187610535517 бф is 0.0 бх is 3.049255642816 97e-05 бц is 1.0164185476056568e-05 бч is 6.776123650704378 5e-06 бш is 6.776123650704378 5e-06 бщ is 0.0003540524607493 0377 бы
is 0.0033778976398761325 бь is 0.0002574926987267 6634 бэ is 0.0001101120093 2394615 бю is 6.776123650704378 5e-06 бя is 0.0007758661580056513
ва is 0.0069675491438367 77 вб is 0.0001778732458 3098992 вв is 0.0003320300588845 1454 вг is 0.0002541046369 0141415 вд is 0.0007199631378873401 ве is 0.0048872791830705 32 вж is
7.961945289577645e-05 вз is 0.0006267914376 90155 ви is 0.0032406811359493 69 вй is 0.0 вк is 0.0009639035893126978 вл is 0.0010164185476056566 вм is 0.0003608285844000 0814 вн is
0.0016262696761690507 во is 0.007331765790062137 вп is 0.0009181647546704432 вр is 0.0009893140530028391 вс is 0.0036794351423324774 вт is 0.0006776123650704378 ву is
0.0008334632090366385 вф is 2.3716432777465325e-05 вх is 9.825379293521348e-05 вц is 7.79254219831 0046e-05 вч is 0.0002557986678140903 вш is 0.0009825379293521349 вщ is
6.7761236507043785e-06 вы is 0.0028713823 9698598 вь is 0.0001355224730 1408757 вэ is 0.0001626269761690508 вю is 3.3880618253521892e-06 вя is 0.0004556943155098694
га is 0.0014314561212113 гб is 5.0820927380282835e-05 гв is 9.825379293521348e-05 гг is 8.470154563380473e-06 гд is 0.0014297620902986238 ге is 0.0001846493694816943 гж is 0.0
гз is 3.049255642816 97e-05 ги is 0.0010638514131605873 гй is 0.0 гк is 0.0001151941020619 7443 гл is 0.0017109712218028555 гм is 3.218658734084579e-05 гн is 0.0004065674190422627
го is 0.0090529011973410 48 гп is 7.623139107042425e-05 гр is 0.0009910080839155152 гс is 8.131348380845254e-05 гт is 6.09851128563394e-05 гу is 0.0007318213542760728 гф is 0.0
гх is 5.082092738028284e-06 гц is 1.6940309126760946e-06 гч is 6.606720559436769e-05 гш is 5.082092738028284e-06 гщ is 1.6940309126760946e-06 гы is 0.0 гь is 0.0 гэ is
1.6940309126760946e-05 гю is 0.0 гя is 1.3552247301408757e-05
да is 0.0046924656281112782 дб is 0.0001067239474985395 дв is 0.001236642566253549 дг is 5.9291081943663306e-05 дд is 9.147766928450 91e-05 де is 0.004719570122715599 дж is
8.639557654648082e-05 дз is 5.251495829295893e-05 ди is 0.0027409422016709921 дй is 0.0 дк is 0.0004235077281690236 дл is 0.0007843361256905708 дм is 0.0001270523184507070 8 дн is
0.0020701057752901876 до is 0.0049177171394987 02 дп is 0.0002862912242422595 др is 0.0012620530299436905 дс is 0.0004844928410253 63 дт is 0.0003540524607493 0377 ду is
0.0019532176423155 37 дф is 3.3880618253521892e-06 дх is 6.945526741971988e-05 дц is 0.0001304403802760593 дч is 7.792542198310034e-05 дш is 0.0001084179784112700 6 дщ is 0.0 ды
is 0.0008029706526084687 дь is 0.0005420898920563503 дэ is 1.1858216388732662e-05 дю is 1.3552247301408757e-05 дя is 0.0005234555520169132
еа is 0.0002134478949971879 еб is 0.0025410463690141417 ев is 0.0034710693400 73318 ег is 0.0044502192076001 ед is 0.0040521219431212 18 ее is 0.0019396653950141283 еж is
0.0010435230422084743 еэ is 0.0023902776177859 69 еи is 0.0010892618768507288 ей is 0.0019227250858873672 ек is 0.0023496208758811743 ел is 0.005720742392107171 ем is
0.0055088885282022 66 ен is 0.0086649681183382 23 ео is 0.0016838667272000 38 еп is 0.0039741965211381 18 ер is 0.0078382810329522 9 ес is 0.0070200641021297356 ет is
0.0102387228362143 15 еу is 0.0008232990235605 82 еф is 9.655976202253739e-05 ех is 0.0009588214965746695 ец is 0.0003811569553521 2126 еч is 0.0023225163812789257 еш is
0.0014822770485915827 ещ is 0.0009063065382817106 еы is 0.0 еь is 0.0 еэ is 0.0003100076570197253 ею is 0.0001439926275774 6803 ея is 0.0005166794283662089
жа is 0.0015754487487887679 жб is 1.0164185476056568e-05 жв is 1.6940309126760946e-05 жг is 2.5410463690141418e-05 жд is 0.0009960901766535436 же is 0.0036150619676507856 жж is
1.3552247301408757e-05 жз is 5.082092738028284e-06 жи is 0.0017567100 5644511 жй is 0.0 жк is 0.0002066717713464 8353 жл is 3.3880618253521892e-05 жм is 4.2350772816902364e-05 жн
is 0.0011587171442704487 жо is 9.655976202253739e-05 жп is 2.3716432777465325e-05 жр is 2.2022401864789228e-05 жс is 1.8634340039437 04e-05 жт is 2.5410463690141418e-05 жу is
0.0004963510574140957 жф is 0.0 жх is 0.0 жц is 0.0 жч is 1.863434003943704e-05 жш is 0.0 жщ is 0.0 жы is 0.0 жь is 1.6940309126760946e-05 жэ is 1.6940309126760946e-06 жю is 0.0
```

таб. 4

H = 7.2301

Для біграм з урахуванням пробілів наведено фрагменти з отриманого файлу after.txt: з перетинами (*таб. 6*) та без (*таб. 5*)

```
110  Bigram for no cross, with spaces
111  аа is 2.2647299451086080e-05 аб is 0.0009993120882791733 ав is 0.0026200094552475206 аг is 0.0005760906797870021 ад is 0.0017848902879887215 ае is 0.0033192448257998034 аж is
     0.0013404370362611574 аз is 0.0032130856096228375 аи is 0.0001910865891185388 ай is 0.0006029843478851668 ак is 0.0053150380899926764 ал is 0.0075231497864076570 ам is
     0.0033475539501136610 ан is 0.0027459850584441873 ао is 6.7941898353258240e-05 ап is 0.0008110564115920202 ар is 0.0018896340479499947 ас is 0.0036419688429777802 ат is
     0.0044997353096876660 ау is 0.0001429610777849808 6 аф is 0.0001075746723926588 8 ах is 0.0011436880622279847 ац is 0.0001104055848240446 4 ач is 0.0013390215800454645 аш is
     0.0008818292223766643 щ is 0.0003269703858250553 аы is 0.0 аь is 0.0 аэ is 1.4154562156928880e-06 аю is 0.0020071169138525040 ая is 0.0021953725905396570 а  is 0.0133576603074937080
112  ба is 0.0012951424373589852 бб is 1.4154562156928799e-05 бв is 3.1140036745243360e-05 бг is 2.1231843235393240e-05 бд is 1.1323649725543040e-05 бе is 0.0017919675690671860 бж is
     9.9081935098501600e-06 бз is 9.9081935098501600e-06 би is 0.0008563510104941924 бй is 0.0 бк is 0.0001288065156280520 8 бл is 0.0006850808083953539 бм is 4.5294598021721600e-05 бн is
     0.0002845066993542688 5 бо is 0.0018287694306752009 бп is 0.0 бр is 0.0010743311267710896 бс is 0.0001981638701970032 бт is 1.4154562156928799e-05 бу is 0.0011252876914758397 бф
     is 0.0 бх is 3.2555492960936240e-05 бц is 7.0772810784643995e-06 бч is 7.0772810784643995e-06 бш is 4.2463686470786400e-06 бщ is 0.0002859221555699617 3 бы is 0.0028040187632875950
     бь is 0.0002349657318050180 7 бэ is 0.0 бю is 4.2463686470786400e-06 бя is 0.0006043998041008597 б  is 0.0002009947826283889 6
113  ва is 0.0057892159221838790 вб is 8.4927372941572800e-06 вв is 4.2463686470786400e-05 вг is 1.2739105941235920e-05 вд is 0.0002406275566677896 ве is 0.0039717701412342210 вж is
     9.9081935098501600e-06 вз is 0.0004373759706490999 3 ви is 0.0024444928845016040 вй is 0.0 вк is 0.0002137338885696248 8 вл is 0.0006935735456895112 вм is 0.0001104055848240446 4 вн
     is 0.0009257083650631435 во is 0.0057651531665171 вп is 0.0001797629393929957 7 вр is 0.0006850808083953539 вс is 0.0025209275201490194 вт is 0.0003085694550210478 ву is
     0.0005576897489829948 вф is 1.4154562156928880e-06 вх is 4.1048230255095352e-05 вц is 2.9724580529555048e-05 вч is 6.7941898353258240e-05 вш is 0.0007374526883759904 вщ is 0.0 вы is
     0.0024034446542465 1 вь is 0.0001019128475298873 6 вэ is 0.0 вю is 4.2463686470786400e-06 вя is 0.0003382940355505983 в  is 0.0054197818498880 38
114  га is 0.0011564277282210828 гб is 0.0 гв is 1.4154562156928799e-05 гг is 1.4154562156928880e-06 гд is 0.0011663359217309333 ге is 0.0001387147091379025 гж is 0.0 гз is
     4.2463686470786400e-06 ги is 0.0008648437477883497 гй is 0.0 гк is 8.7758285372958560e-05 гл is 0.0014310262340655017 гм is 8.4927372941572800e-06 гн is 0.0002618593999031828 го is
     0.0075330579799175075 гп is 2.8309124313857600e-06 гр is 0.0008563510104941924 гс is 1.6985474588314560e-05 гт is 4.3879142686479280e-05 гу is 0.0006143079976107099 гф is 0.0 гх is
     0.0 гц is 0.0 гч is 3.9632774039400640e-05 гш is 2.8309124313857600e-06 гщ is 0.0 гы is 0.0 гь is 0.0 гэ is 0.0 гю is 0.0 гя is 0.0 г  is 0.0006100616289636312
115  да is 0.0039264755423320490 дб is 7.2188267000336870e-05 дв is 0.0009625102266711584 дг is 1.6985474588314560e-05 дд is 2.6893668098164720e-05 де is 0.0039675237725871420 дж is
     9.3420110235730070e-05 дз is 1.4154562156928799e-05 ди is 0.0022321744521476717 дй is 0.0 дк is 0.0002250575382951679 дл is 0.0006100616289636312 дм is 4.8125511333557920e-05 дн
     is 0.0015357699940267747 до is 0.0039956200122861599 дп is 0.0001231446907652805 6 др is 0.0011083022168875251 дс is 0.0002576130125610417 дт is 0.0002179802572167035 3 ду is
     0.0017169483896354635 дф is 0.0 дх is 4.5294598902172160e-05 дц is 0.0001004973913141944 8 дч is 3.1140036745243360e-05 дш is 6.9357354568951120e-05 дщ is 0.0 дь is
     0.0007332063197289118 дь is 0.0004203904960760853 6 дэ is 1.4154562156928880e-06 дю is 9.9081935098501600e-06 дя is 0.0004317141457863284 д  is 0.0012201232579272625
116  еа is 5.3787336196329440e-05 еб is 0.0013376061238297716 ев is 0.001331944298967 ег is 0.0033801094430745975 ед is 0.0025350820823059488 ее is 0.0013206206492414570 еж is
     0.0007289599510818331 ез is 0.0012441860135904150 4 еи is 7.6434635647415520e-05 ей is 0.0015853109615760255 ек is 0.0011394422536327683 ел is 0.0045535226458839950 ем is
     0.0039491228417831350 ен is 0.0055259410660650040 ео is 0.0001910865891185388 еп is 0.0014550889897322805 ер is 0.0061954518560877360 ес is 0.0037651135337430607 ет is
     0.0078911684024878060 еу is 0.0001118210410397375 1 еф is 1.1323649725543040e-05 ех is 0.0005675979424928448 ец is 0.0002915839804327325 5 еч is 0.0013818485266516251 еш is
     0.0011606740968681615 ещ is 0.0007606999878270765 еы is 0.0 еь is 0.0 еэ is 1.4154562156928880e-06 ею is 0.0001203137783338948 ея is 0.0002236420820794750 2 е  is
     0.0176748016765356993
117  жа is 0.0013984707411045655 жб is 8.4927372941572800e-06 жв is 2.8309124313857600e-06 жг is 2.1231843235393240e-05 жд is 0.0007374526883759904 же is 0.0030885254626418640 жж is
     5.6618248627715260e-06 жз is 0.0 жи is 0.0014239489529870373 жй is 0.0 жк is 0.0001528692712948310 3 жл is 2.1231843235393240e-05 жм is 2.4062755666778960e-05 жн is
     0.0008761673975138927 жо is 6.5110985921872480e-05 жп is 1.4154562156928880e-06 жр is 1.9816387019700320e-05 жс is 2.8309124313857600e-06 жт is 0.0 жу is 0.0003821731782370776 жф is
```

таб. 6

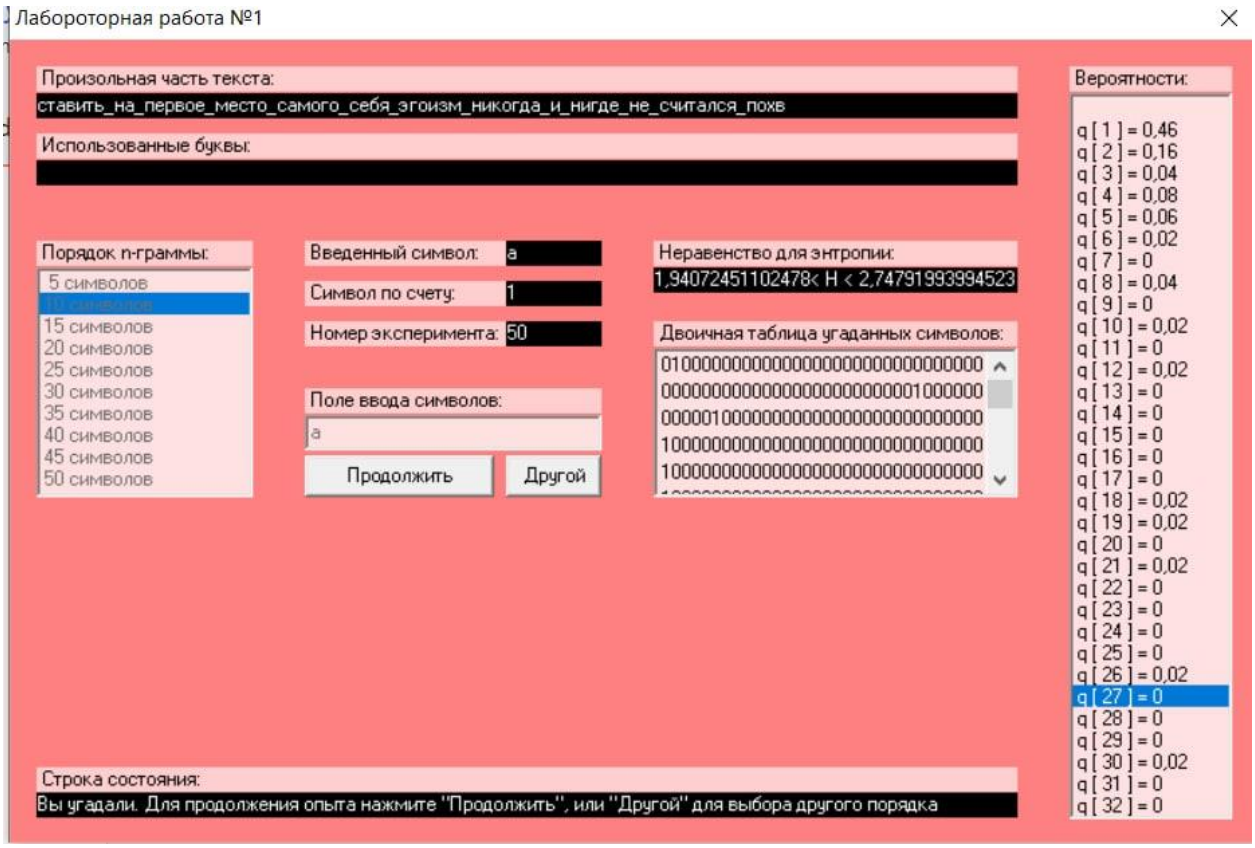
H = 36.3867

```
76   Bigram cross, with spaces
77   аа is 2.3355027558932518e-05 аб is 0.0010177130190831807 ав is 0.0026001930682278203 аг is 0.0005973225230223953 ад is 0.0017962139377142646 ае is 0.0033390612128195037 аж is
     0.0013163742805943784 аз is 0.0031621291858578940 аи is 0.0001868402204714601 5 ай is 0.0006093539008557848 ак is 0.0052414343667107340 ал is 0.0075404829891038180 ам is
     0.0033461384938979683 ан is 0.0027141372935910973 ао is 6.1572345382640280e-05 ап is 0.0008245032456411026 ар is 0.0019243127252344704 ас is 0.0036193215435266940 ат is
     0.0045096435031975160 ау is 0.0001323451561672842 8 аф is 0.0001146519534711223 7 ах is 0.0011153794979659895 ац is 0.0001089901286083517 6 ач is 0.0013234515616728428 аш is
     0.0009065997061512896 щ is 0.0003425404041976769 7 ы is 0.0 аь is 0.0 аэ is 2.1231843235393240e-06 аю is 0.0020333028538428223 ая is 0.0021564475446081027 а  is
     0.0133739380539741760
78   ба is 0.0013008042622217567 бб is 1.3446834049082360e-05 бв is 2.6893668098164720e-05 бг is 1.6277746480468120e-05 бд is 1.0615921617696660e-05 бе is 0.0018266462463516661 6 бж is
     1.2031377833389480e-05 бз is 1.3446834049082360e-05 би is 0.0008747519412981999 6 й is 0.0 бк is 0.0001358837967065165 бл is 0.0006921580894738183 бм is 4.1048230255095352e-05 бн is
     0.0002661057685502614 5 бо is 0.0018549553706655193 бп is 0.0 бр is 0.0010191284752988736 бс is 0.0009179431722638523 бт is 8.4927372941572800e-06 бу is 0.0011104254012110643 бф
     is 0.0 бх is 2.1231843235393240e-05 бц is 6.3695529706179600e-06 бч is 5.6618248627715260e-06 бш is 6.3695529706179600e-06 бщ is 0.0003099849112367407 бы is 0.0028195888716602170 бь is
     0.0002179802572167035 3 бэ is 0.0 бю is 4.9540967549250800e-06 бя is 0.0006220930067970207 б  is 0.0002137338885696248 8
79   ва is 0.0057715227194877180 вб is 9.2004654020037260e-06 вв is 4.0340502147247080e-05 вг is 1.4154562156928799e-05 вд is 0.0002385043723442502 7 ве is 0.0040191879249459933 вж is
     7.0772810784643995e-06 вз is 0.0004508228046981822 5 ви is 0.0024989879488057794 вй is 0.0 вк is 0.0002215188977559357 3 вл is 0.0007183440294641366 вм is 0.0001012051194220409 2
     вн is 0.0009158001715532933 во is 0.0058698969264783735 вп is 0.0001861324923636137 вр is 0.0006949890019052041 вс is 0.0025718839439139630 вт is 0.0003099849112367407 ву is
     0.0005683056706006913 вф is 7.0772810784644e-07 вх is 3.9632774039400640e-05 вц is 2.9724580529555048e-05 вч is 7.3603723216029750e-05 вш is 0.0007381604164838369 вщ is
     7.0772810784644e-07 вы is 0.0023963673731680460 вь is 0.0001068669442848412 44 вэ is 0.0 вю is 2.1231843235393240e-06 вя is 0.0003290935701485946 в  is 0.0053560863201818580
80   га is 0.0012045532395546408 гб is 7.0772810784644e-07 гв is 1.7693202696160998e-05 гг is 1.4154562156928880e-06 гд is 0.0011543045438975436 ге is 0.0001344683404908236 гж is 0.0
     гз is 4.9540967549250800e-06 ги is 0.0008315805267195670 гй is 0.0 гк is 8.2096460510187040e-05 гл is 0.0014097943908301085 гм is 1.2031377833389480e-05 гн is 0.0002576130312561041 7
     го is 0.0075160725053291930 гп is 1.4154562156928880e-06 гр is 0.0008492737292415728 гс is 1.4154562156928799e-05 гт is 4.0340502147247080e-05 гу is 0.0005909529700517774 гф is 0.0
     гх is 0.0 гц is 0.0 гч is 4.9540967549250800e-05 гш is 3.5386405392321998e-06 гщ is 0.0 гы is 0.0 гь is 0.0 гэ is 0.0 гю is 0.0 гя is 7.0772810784644e-07 г  is
     0.0005711365830320770
81   да is 0.0038726882061357196 дб is 6.8649626461104680e-05 дв is 0.0009398629272200723 дг is 1.2739105941235920e-05 дд is 2.5478211882471840e-05 де is 0.0039625696758322170 дж is
     7.2188267000336870e-05 дз is 9.9081935098501600e-06 ди is 0.0022350053645790575 дй is 0.0 дк is 0.0002356734599128645 дл is 0.0006079384446400919 дм is 5.5202792412022320e-05 дн is
     0.0014862290264775239 до is 0.0039639851320479105 дп is 0.0001415456215692879 8 др is 0.0010700848990638173 дс is 0.0002491202939619469 дт is 0.0002278884507265536 7 ду is
     0.0016447601226351265 дф is 0.0 дх is 5.6618248627715196e-05 дц is 0.0001104055848240446 4 дч is 2.4062755666778960e-05 дш is 8.1388732402340 6e-05 дщ is 7.0772810784644e-07 ды is
     0.0007254213105426010 дь is 0.0004678082792864968 дэ is 7.0772810784644e-07 дю is 1.1323649725543040e-05 дя is 0.0004232214084921711 д  is 0.0012427705573783486
82   еа is 5.0956423764943680e-05 еб is 0.0013149588243786855 ев is 0.0013453911330160823 ег is 0.0033666626090255515 ед is 0.0025669298471590377 ее is 0.0013338313851937618 еж is
     0.0007183440294641366 ез is 0.0012087996082017194 еи is 7.4311451323876 2e-05 ей is 0.0016150355421055760 ек is 0.0011684591060544724 ел is 0.0045683849361487705 ем is
     0.0039646928601557570 ен is 0.0054771078266235990 ео is 0.0001762242988537635 7 еп is 0.0014621662708107450 ер is 0.0059930416172436540 ес is 0.0037268962159193530 ет is
     0.0079757694844625371 еу is 0.0001104055848240446 4 еф is 1.5570018372621680 5e-05 ех is 0.0005654747581693055 ец is 0.0002901685242170403 7 еч is 0.0014260721373105766 еш is
     0.0011557200001132365 ещ is 0.0007671772689055409 еы is 0.0 еь is 0.0 еэ is 2.1231843235393240e-06 ею is 0.0001160674096868161 5 ея is 0.0002285961788344001 е  is
     0.0175799661989055680
83   жа is 0.0013383133851937618 жб is 7.0772810784643995e-06 жв is 4.2463686470786400e-06 жг is 2.2647299451086080e-05 жд is 0.0007735468218761589 же is 0.0030276608453670703 жж is
     1.0615921617696660e-05 жз is 0.0 жи is 0.0014317339621733480 жй is 0.0 жк is 0.0001535769994026774 7 жл is 2.2647299451086080e-05 жм is 2.3355027558932518e-05 жн is
```

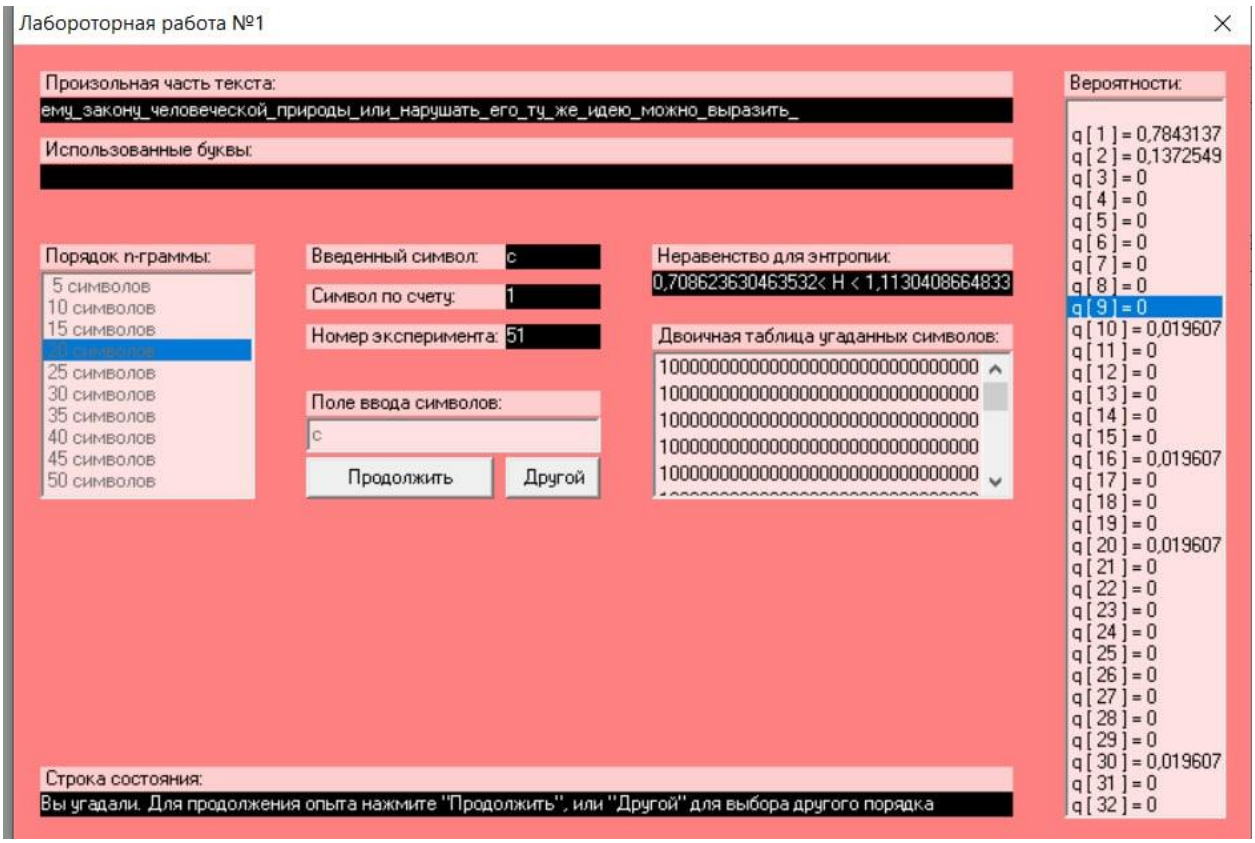
таб. 4

H = 36.406

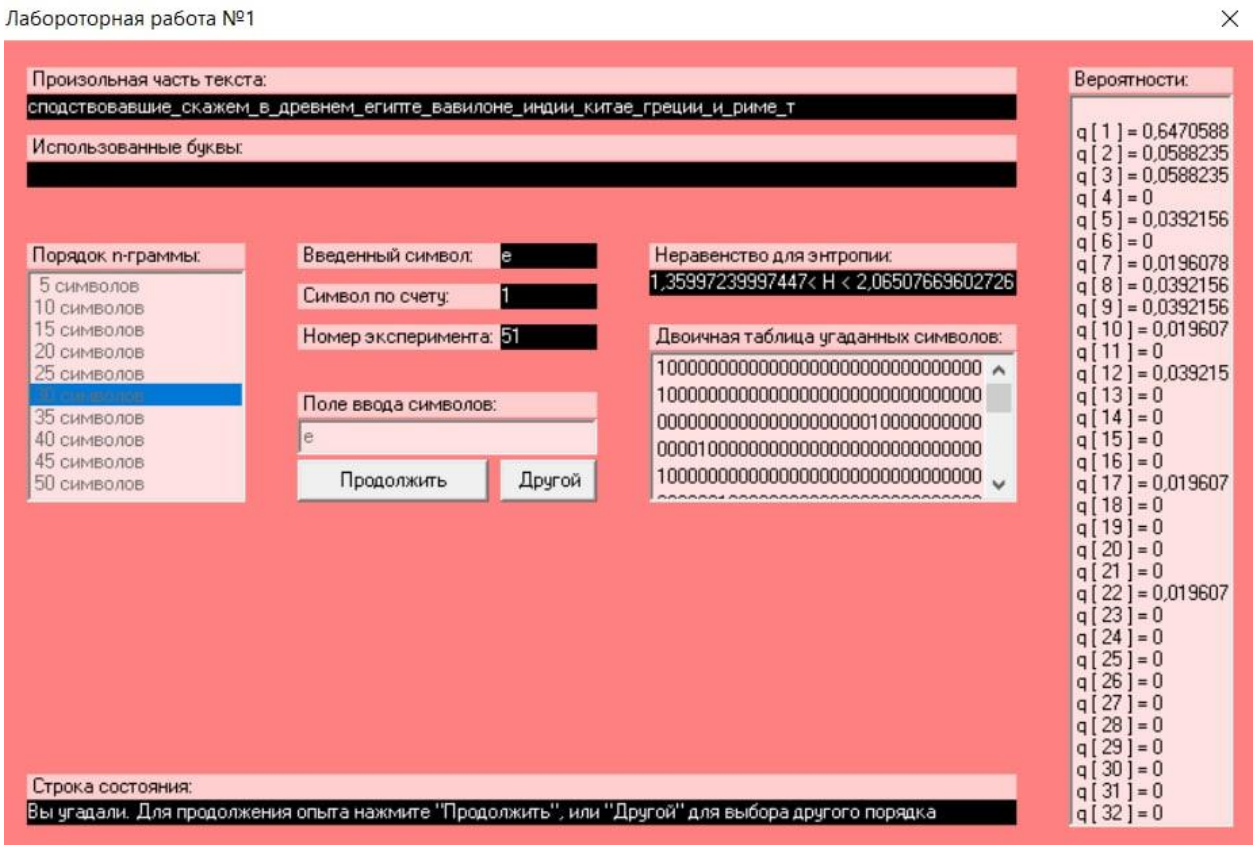
За допомогою CoolPinkProgram було оцінено значення H^{10}, H^{20}, H^{30}.



$$H^{10} = \frac{1,94+2,75}{2} = 2,345$$



$$H^{20} = \frac{0,71+1,11}{2} = 0,91$$



$$H^{30} = \frac{1,36+2,07}{2} = 1,715$$

H_0 для заданого алфавіту – $\log_2 32=5$, тоді:

1) $R = 1 - \frac{2,345}{5} = 0,531$
2) $R = 1 - \frac{0,91}{5} = 0,818$
3) $R = 1 - \frac{1,715}{5} = 0,65$

Середнє значення R надлишковості російської мови – 0,666, але кількість проведених експериментів ніяк не прямує до нескінченності, тому можуть виникати значні похибки.

Висновки: після виконання цієї лабораторної роботи в нас з’явилося більш чітке уявлення про поняття ентропії не лише у вигляді формули, а на конкретному прикладі мови. Підтвердили, що не має істотної різниці у розподілі частот при підрахунку їх для біграм з перетином і без, що свідчить про коректність програми.

І з першої, і з другої частини роботи побачили, як часто зустрічається пробіл. Друга частина показала, що надлишковість російської мови – приблизно 66%, що свідчить про повторюваність значної кількості інформації, але є цілком природнім для усіх мов.

Також було отримано навички роботи з мовою Python та системою Git.